REVENUE MANAGEMENT FOR HOTEL BOOKINGS

Submitted by:

Group 11

Amey Ghate (ag84957)

Sameer Ahmed (sa49847)

Spandan Pal (sp54969)

Submitted to:

Dr. Guoming Lai

Professor

May 2024

INTRODUCTION

The hotel industry is a dynamic and competitive sector where the effective management of room bookings and pricing strategies can significantly impact revenue generation and market positioning. Within this context, the concept of revenue management becomes crucial, as it allows hotel managers to optimize the balance between room occupancy and room rates, ultimately enhancing profitability. This paper explores the application of sophisticated revenue management techniques to a 100-room hotel, focusing on the strategic allocation of rooms to different guest segments to maximize revenue.

Traditionally, hotels have relied on securing large group bookings to ensure high occupancy rates. While such bookings can guarantee a steady stream of revenue, they often come at the expense of potentially higher-paying transient guests. Transient guests, who book individually and often last-minute, might pay higher rates per room but pose a higher risk of unsold inventory if not managed correctly. The challenge for hotel managers is to find an optimal balance that maximizes occupancy and average daily rate (ADR) without sacrificing potential higher revenues from transient guests.

This paper presents an analytical approach to address this challenge, employing statistical models and data analysis to guide decision-making in room allocation and pricing. By integrating principles from operations research and economics, particularly capacity control and pricing models, the analysis provides insights into how hotels can enhance their revenue management strategies effectively.

PROBLEM STATEMENT

The primary challenge addressed in this study revolves around the decision-making process concerning room allocation between group and transient bookings in a 100-room hotel. The hotel management must determine how many rooms to reserve for higher-paying transient guests versus offering them to groups at potentially lower rates. This dilemma is exacerbated by the inherent uncertainty of transient guest bookings and the financial stability that group bookings can offer.

Specifically, the problem can be framed as follows: How should the hotel allocate its rooms to balance the immediate financial security provided by group bookings with the potential for higher revenue from transient guests who may book closer to their arrival date?

The decision is complicated by various factors including the seasonal demand patterns, the average length of stay, and the differing price sensitivities of these two market segments.

The aim is to develop a decision-support model that can dynamically adjust room allocations in real-time, based on current booking trends, historical data, and forecasted market conditions. Such a model would allow the hotel to enhance its revenue per available room (RevPAR), a critical metric in hotel revenue management.

DATASET OVERVIEW

As a team of hotel analysts and data scientists, we chose the "Hotel Booking Demand" dataset (https://www.kaggle.com/datasets/jessemostipak/hotel-booking-demand) to delve into the complex dynamics of hotel booking practices. This dataset provides an exceptional foundation for our analysis, offering comprehensive booking information for two distinct types of accommodations: a city hotel and a resort hotel. Our goal is to harness this data to enhance our understanding of the factors influencing hotel bookings and to refine our skills in data analysis and predictive modeling.

Context

Our interest in this dataset stems from its potential to answer pivotal questions that are crucial for anyone involved in the hospitality industry. For instance, determining the best times of year to book a hotel room, the optimal length of stay to ensure the best daily rates, and predicting the likelihood of guests making special requests. These insights are invaluable for optimizing booking strategies and enhancing guest satisfaction, areas we are keen to explore and understand deeply.

	hotel	is_canceled	lead_time	arrival_date_year	arrival_date_month	arrival_date_week_number	arrival_date_day_c
0	Resort Hotel	0	342	2015	July	27	1
1	Resort Hotel	0	737	2015	July	27	1
2	Resort Hotel	0	7	2015	July	27	1
3	Resort Hotel	0	13	2015	July	27	1
4	Resort Hotel	0	14	2015	July	27	1

Content

The dataset encompasses a wealth of information that is vital for our analysis:

- Booking Timeframe: It includes data on when bookings are made, which helps us analyze booking lead times and their impact on the success of bookings.
- Stay Details: It documents the length of stay and the number of adults, children, and babies per booking, providing insights into guest demographics and preferences.

- Facilities Usage: Information on the use of additional facilities like parking spaces offers insights into the ancillary services that may influence guest choices.
- Room Rates: The inclusion of daily rates allows us to analyze pricing strategies across different times of the year.
- Anonymity: All personally identifying information has been removed, focusing our analysis on behavioral and transactional data.

This dataset comprises 119,390 rows and 32 columns, each providing detailed attributes of bookings made at the city and resort hotels.

Data Utilization

One pivotal column in our analysis is the is_canceled column, which indicates whether a booking was canceled before the guest's arrival. By filtering out the entries where is_canceled equals 1, we ensured our analysis was concentrated only on those bookings that actually resulted in stays.

This approach allowed us to accurately measure and evaluate the real occupancy and revenue, avoiding skewness in data interpretation and ensuring our findings reflected genuine customer behavior and hotel performance.

Data Source and Preparation

Originally published in the scholarly article "Hotel Booking Demand Datasets" by Nuno Antonio, Ana Almeida, and Luis Nunes, this data was later cleaned and made accessible by Thomas Mock and Antoine Bichat for the #TidyTuesday project. This initiative not only enhanced the dataset's usability but also aligned with our commitment to ethical data use and privacy.

Inspiration and Use

This dataset was a perfect match for our objectives, offering both a challenge and an opportunity to practice and improve our data analysis and modeling skills. Whether we were exploring data visualizations to understand consumer behavior or developing machine learning models to forecast booking trends, this dataset provided a comprehensive base for our work.

The practical applications of this dataset extended beyond academic exercises. By analyzing this data, we aimed to produce actionable insights that could be used by hotel managers to adjust pricing, anticipate peak booking periods, and tailor marketing strategies to better target specific guest segments. Moreover, it served as a critical educational tool for us, enhancing our ability to work with real-world data in the hospitality industry context.

In summary, our choice to work with the "Hotel Booking Demand " dataset was driven by our desire to blend theoretical knowledge with practical application, ensuring we not only understood the data but also contribute meaningfully to the field of hospitality management. Through this project, we developed robust analytical skills that will underpin our future careers as data scientists and analysts.

CAPACITY CONTROL MODEL

Within the domain of revenue management in the hospitality industry, capacity control models play a pivotal role by optimizing the allocation and pricing of hotel rooms based on anticipated demand.

Capacity control refers to the techniques and strategies employed to manage the availability of resources—in this case, hotel rooms—to achieve the best financial outcome. It integrates both inventory management and pricing strategies to balance demand with supply efficiently.

Hotels typically utilize various models of capacity control, each with its unique approach to managing room inventory and pricing:

- **Fixed Allocation Model:** This traditional model involves allocating a predetermined number of rooms to different market segments, such as business travelers, tour groups, and leisure travelers. The allocation is based on historical data and does not change in response to current market conditions.
- **Dynamic Allocation Model**: More advanced than its fixed counterpart, the dynamic allocation model allows for real-time adjustments in room allocation and pricing. It leverages live data streams and demand forecasting techniques to optimize revenue. For instance, prices may increase as availability decreases, or special promotions might be offered during periods of low demand.

In this analysis, we will use the Fixed Allocation Model. To reduce the complexity of the problem, we will consider only two classes right now.

- Group Customers : refer to individuals who make reservations as part of a larger group. These might include *corporate groups, social groups or tour groups*. They are typically negotiated at a fixed rate and may include specific contractual agreements regarding room blocks, event spaces, and other services. The main characteristics of group customers include bulk booking, planned well in advance, and often linked to event-based stays.
- Transient Customers : refer to individual travelers or small groups who book rooms independently, not as part of a contracted or organized group. This segment includes *business travelers, leisure travelers, and walk-in guests*. They are more variable and unpredictable than group bookings, influenced heavily by seasonal trends, pricing strategies, and individual preferences.

In the Capacity Control Model implemented, we compare the Expected Revenue from two outcomes (i.e securing the room for Group customers vs reserving the room for Transient customers) and choose the outcome for which the Expected Revenue is greater.

$$Expected\ Revenue\ =\ Outcome\ Revenue\ \times\ Probability\ of\ Outcome$$

Where:

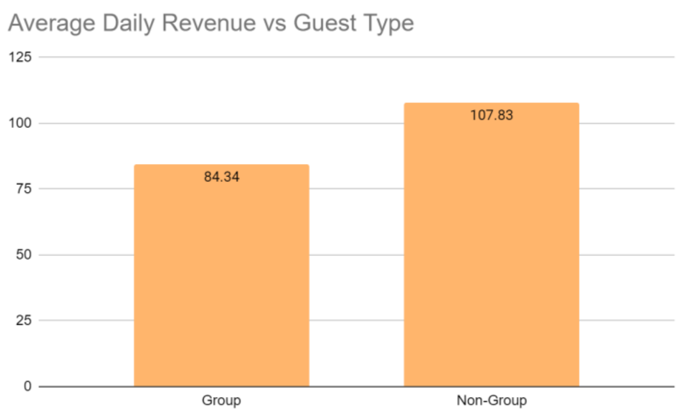
- If $Expected\ Revenue\ (group)\ <\ Expected\ Revenue\ (transient)$: We hold the rooms for the High Value Transient Guests.
- Else $Expected\ Revenue\ (group)\ >\ Expected\ Revenue\ (transient)$: We allocate the room to the Group Customer.

The probability of arrival of transient customers is unknown currently. This probability affects the expected revenue from transient customers and thus the amount of rooms to hold.

Now, we make certain assumptions to simplify our analysis :

1) All bookings are set to have a Length Of Stay of 1 day
2) Canceled or No-Show Records are not being considered.

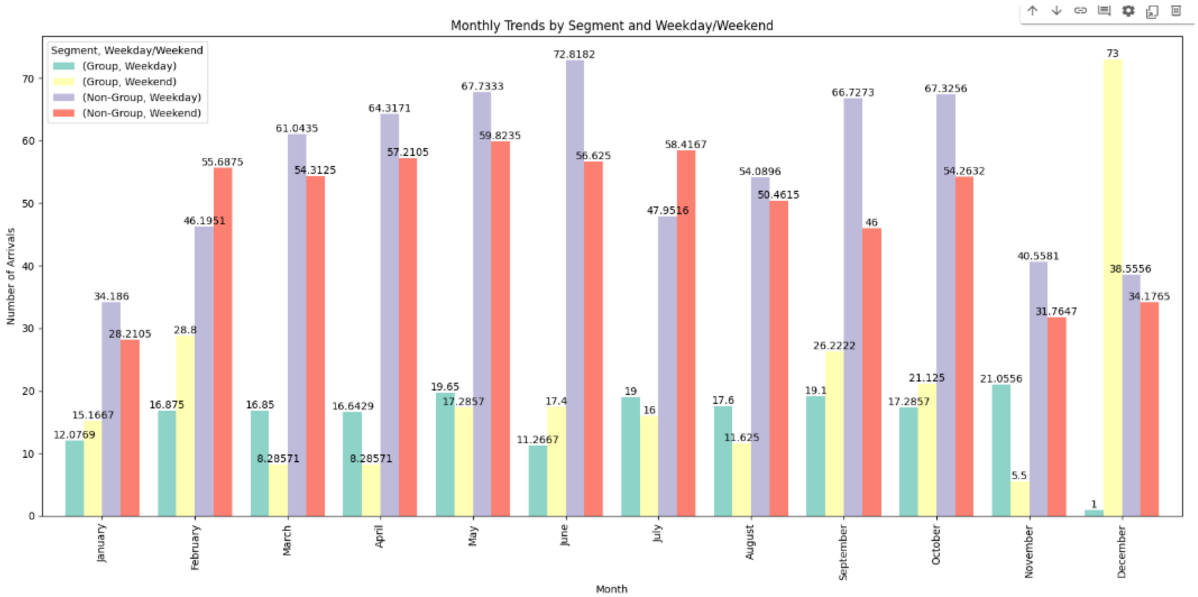
Looking at the data, we see that Group Customers have a lower average daily rates compared to Transient Customers. This indicates alignment with the capacity control model.



DEMAND ANALYSIS

The task here is to estimate the probability of arrival of transient customers. Other than segmentation on the basis of type of customer (i.e. group vs transient), the arrival is also segmented on the basis of the type of day (i.e. weekday or weekend)

We get the following plot if we plot the historical data.



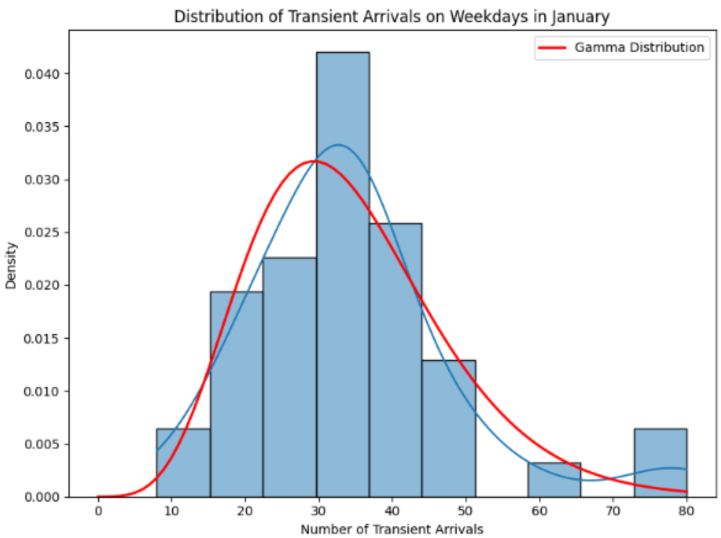
The following can be noted from the customer arrival distribution.

1) Transient Bookings:
    - Transient bookings dominate across all months, both on weekdays and weekends.
    - The highest number of transient room nights sold occurs in July, and generally, the summer months (June, July, August) show peak activity.
2) Group Bookings:
    - Group bookings are generally lower than transient but show a notable peak in December, which is the highest point for group bookings in the year.
    - There is also a smaller peak in September for group bookings.
3) Weekday vs. Weekend:
    - For all segments, the number of room nights sold tends to be higher on weekdays than weekends, with a few exceptions, notably in the group segment in May and October and the transient segment in October.

4) Seasonal Trends:
    - There is a clear seasonal trend with higher bookings during the middle of the year (summer) and another smaller peak in the fall, particularly noticeable in October.
    - The winter months (January, February) and late fall (November, December) generally show the lowest room nights sold across all segments.

Taking this into account, we try to estimate the probability distribution for the arrival of transient customers.

For example, we take the distribution of arrival of transient customers for Weekdays in January and then we try to obtain a function that fits this distribution in a satisfactory fashion. Initially, we tried a Normal distribution fit the data well but it does not fit the data well. But eventually, it was observed that the Gamma distribution does track the data to a great extent.



Now, for the probability distribution function for a Gamma distribution, we needed to estimate three parameters.

$$f(x; k, \theta) = \frac{x^{k-1} e^{-x/\theta}}{\theta^k \Gamma(k)} \quad \text{for } x > 0 \text{ and } k, \theta > 0.$$

Here, x, k and θ have specific roles:

- x (Variable):

  This is the variable of the distribution, representing the data or outcome values over which the distribution is defined. The condition

  - x>0
  - x>0 indicates that the Gamma distribution is defined only for positive values.

2. k (Shape parameter, sometimes denoted as α):

   This parameter influences the shape of the distribution. It is often referred to as the shape parameter. Depending on the value of k, the distribution can be highly skewed (for small k) or approach a normal distribution as k increases. The requirement that k>0 ensures the mathematical validity of the distribution (i.e., the integrals and calculations involved yield real, positive numbers).

3. θ (Scale parameter, sometimes denoted as $\beta$where β=1/θ):
   This parameter sets the scale of the distribution. It stretches or compresses the distribution along the x-axis. Higher values of θ lead to a wider, more spread-out distribution, while smaller values lead to a narrower one. Similar to k, θ must be positive.

Considering this, if we proceed to estimate the parameters for Gamma probability distribution functions for each of the 24 Month + Weekday/Weekend combinations, we obtain the following.

Month	Day Type	x	k	theta
July	Weekday	304.3417	-582.055	1.758321
July	Weekend	1099.499	-894.928	0.86706
August	Weekday	465.6156	-590.438	1.090953
August	Weekend	411.6385	-379.301	1.052574
September	Weekday	29.06477	-66.0485	4.562873
September	Weekend	4.253622	-6.62511	1.504757
October	Weekday	4.62693	5.928041	13.26595
October	Weekend	4.28277	0.149799	12.63514
November	Weekday	0.7995	7	27.93783
November	Weekend	1.155852	10.86138	35.69747
December	Weekday	2.128556	-3.68894	16.3781
December	Weekend	48.1758	-82.7169	2.42456
January	Weekday	7.516583	-22.1512	4.848083
January	Weekend	2.734694	13.47417	2.692006
February	Weekday	3.733156	9.762075	9.75924
February	Weekend	0.562418	34	22.50036
March	Weekday	26.96081	-4.8323	2.436954
March	Weekend	2.168334	31.52253	10.48038
April	Weekday	9.543699	5.123834	5.208372
April	Weekend	8.567586	-31.916	12.5182
May	Weekday	100.0366	-49.1722	0.542047
May	Weekend	29.789	-42.2512	3.147205
June	Weekday	27.52021	33.34962	14.47896
June	Weekend	12.45411	-51.3598	7.575966

OPTIMAL POLICY

As mentioned before, Expected revenue from a booking class can be calculated as:

$$Expected\ Revenue\ = r \times p$$

where,

r is the rate charged per room,

p is the probability of selling the room.

This formula provides a straightforward metric for assessing the potential revenue from different customer segments or booking classes under static conditions. However, the dynamic nature of hotel bookings, where the decision to accept a booking today could preclude higher revenue from a future booking, necessitates a more nuanced approach.

This is where **Littlewood's Rule** comes into play, bridging the gap between theoretical revenue management and practical, decision-driven strategies. Littlewood's Rule, a fundamental principle in revenue management, is designed to optimize the decision-making process regarding whether to accept a current booking request or wait for a potentially higher-paying customer. The rule is particularly useful in scenarios with multiple fare classes and limited capacity.

Littlewood's Rule states that a lower fare class booking should be accepted if and only if the expected revenue from accepting the booking exceeds the expected marginal revenue from reserving capacity for a higher fare class. Mathematically, this can be expressed as:

$$r \ \geq \ p_H \ \times \ (r_H - r)$$

where,

r is the revenue from the current lower fare class.

r_H is the revenue from a potential future higher fare class.

p_H is the probability of demand for the higher fare class arriving after accepting the current booking.

Here, $p_H \times (r_H - r)$ represents the expected opportunity cost of accepting the lower fare class now — essentially, what is potentially lost in revenue if a higher paying customer arrives later. The rule succinctly encapsulates the trade-off between immediate and future gains, guiding managers in setting booking limits that optimize overall revenue.

This can be simplified to a general rule,

$$If\ P(selling\ more\ than\ X\ High\ Value\ guests) > \frac{Revenue\ From\ Group}{Revenue\ From\ Non - Group}$$

Then, we set aside X number of rooms for High Value guests.

Now, to find out X here, we do the following steps

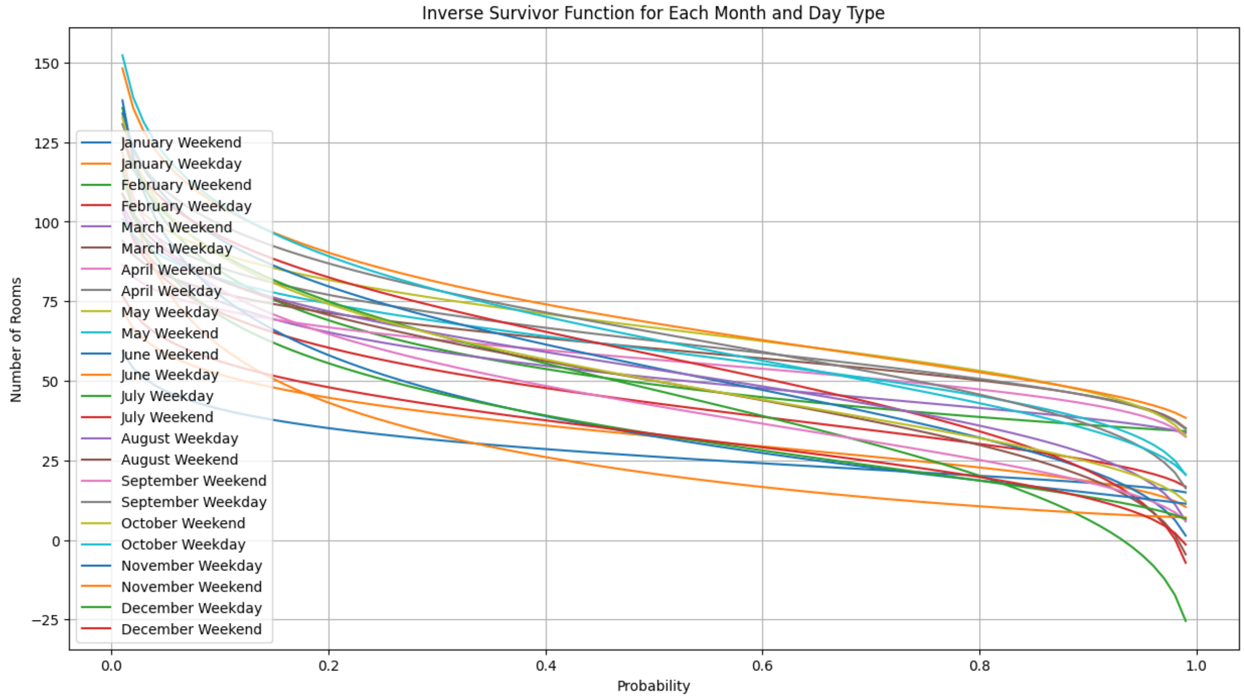
- Calculate the ratio of revenue from group customers to revenue from transient customers
- Use the Inverse Survivor Function on the ratio obtained in the previous step to find X i.e. the number of rooms to be set aside for transient occupancy.

RESULTS

Implementing the optimal policy based on our refined understanding of Littlewood's Rule and expected revenue calculations has significantly enhanced our revenue management strategies. This strategic implementation involves setting aside a calculated number of rooms for high-paying transient guests, using a dynamic decision-making framework.

By applying the established rule: $If\ P(selling\ more\ than\ X\ High\ Value\ guests) > \frac{Revenue\ From\ Group}{Revenue\ From\ Non-Group}$, we actively adjusted our room allocation across different booking classes. To determine the optimal number of rooms X to reserve, we calculated the revenue ratio between group customers and transient customers. Further, we utilized the Inverse Survivor Function on the calculated ratio to accurately find X.

The provided graph illustrates the Inverse Survivor Function for various guest segments categorized by month and day type (weekday or weekend). Each curve represents a different combination of month and day, highlighting the distinct booking patterns and the probability of room occupancy across these segments.



The provided graph illustrates the Inverse Survivor Function for various guest segments categorized by month and day type (weekday or weekend). Each curve represents a different combination of month and day, highlighting the distinct booking patterns and the probability of room occupancy across these segments.

From the graph, we observe that the Inverse Survivor Functions vary significantly between different months and whether the days are weekdays or weekends. For example, the steeper curves for certain months on weekends might indicate a higher likelihood of room occupancy compared to weekdays, suggesting a greater demand during these times. This variance allows us to effectively segment our guest population based on their booking behaviors and the probability of room bookings.

Key observations include:

- Months like June and July show higher curves on weekends, which could indicate peak tourist season leading to increased bookings.
- The negative slope observed in December for weekdays suggests a lower probability of room occupancy, possibly due to reduced travel during the holiday season.
- The diversity in the slopes and intersections of these functions confirms that each segment—defined by month and day type—has unique characteristics in terms of room demand.

This segmentation based on the Inverse Survivor Function enables us to tailor our room pricing and availability strategies more effectively; creating an optimal policy.

This policy has been applied throughout the year, and the results show a strategic allocation of rooms that optimally balances between immediate gains from group bookings and potential higher revenues from transient guests. Here is a detailed summary of the number of rooms reserved for transient guests, differentiated by weekdays and weekends throughout the months:

Month	Weekday/Weekend	No of Rooms Reserved For Transient Guests
January	Weekday	10
January	Weekend	18
February	Weekday	29
February	Weekend	39
March	Weekday	39
March	Weekend	33
April	Weekday	43
April	Weekend	50
May	Weekday	52
May	Weekend	49
June	Weekday	41
June	Weekend	56
July	Weekday	30
July	Weekend	48
August	Weekday	45
August	Weekend	41
September	Weekday	26
September	Weekend	44
October	Weekday	36
October	Weekend	40
November	Weekday	20
November	Weekend	10
December	Weekday	0
December	Weekend	23

From this we are now able to calculate the expected profit to see the impact of our model. To calculate the expected profit with a capacity control model using gamma distribution parameters, we go through the following steps:

1) Model Room Demand: We use the gamma distribution to model the number of rooms that are expected to be booked. The expected number of rooms sold is calculated as the mean of the gamma distribution for each month and day type (weekday or weekend).
2) Compute Revenue: We calculate the revenue by using the average daily rates for groups and transient bookings. The formula used is:

$$\text{Adjusted ADR} = (\text{Group ADR} + \text{Transient ADR}) \times (1 + \frac{\text{Transient Protection}}{100})$$

3) Expected Profit Calculation: The expected profit is then computed by multiplying the expected number of rooms sold by the adjusted average daily rate (ADR). This gives us the profit considering the capacity control model:

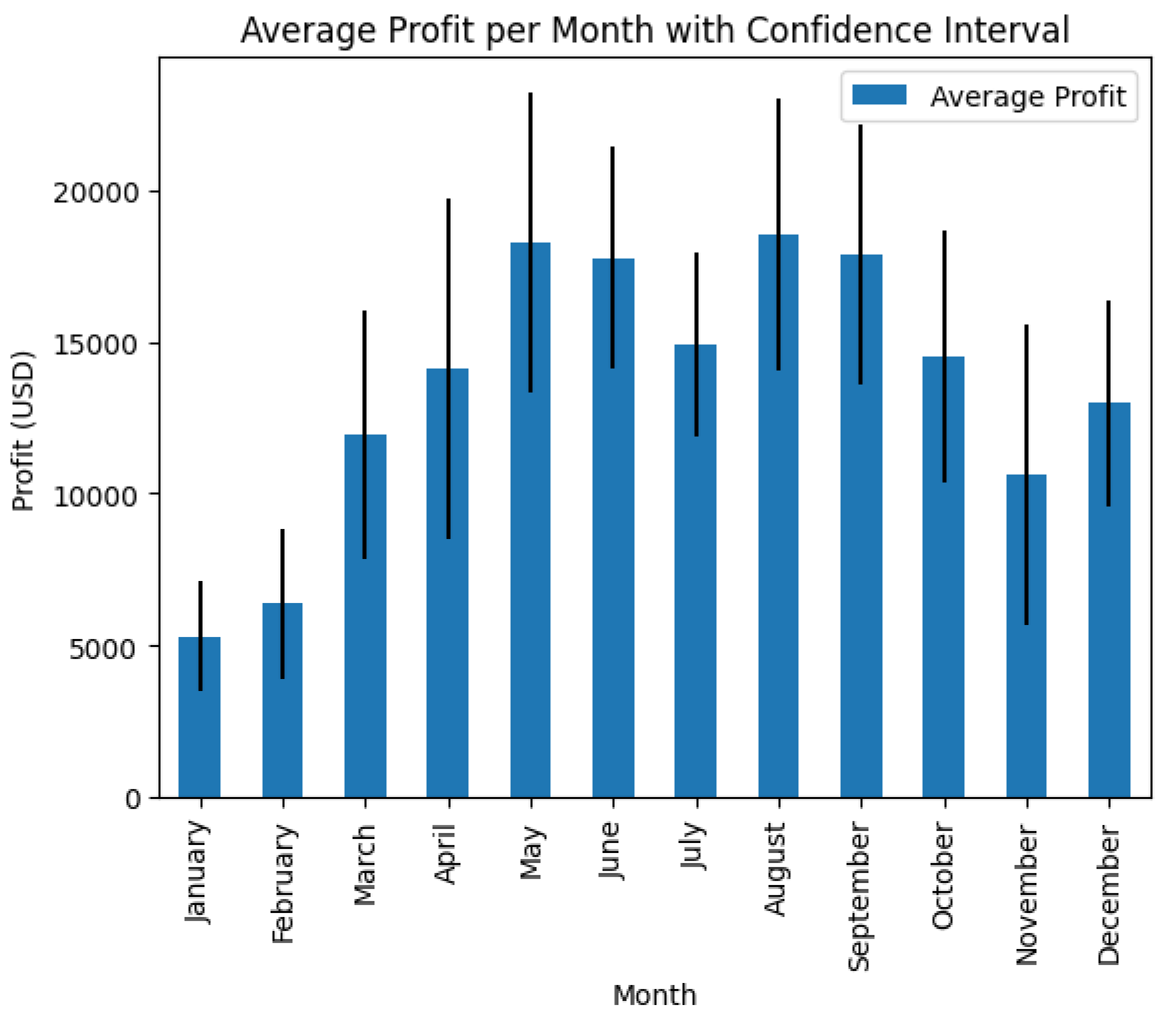
$$\text{Expected Profit} = \text{Expected Rooms Sold} \times \text{Adjusted ADR}$$

Month	**Day Type**	**Expected Profit (Model)**	**Expected Profit (No Model)**	**Difference**
January	Weekend	5326.15	7500	-2173.85
January	Weekday	6358.61	7200	-841.39
February	Weekend	12154.95	6900	5254.95
February	Weekday	9594.27	7300	2294.27
March	Weekend	13060.53	8100	4960.53
March	Weekday	14517.88	8900	5617.88
April	Weekend	17585.74	10100	7485.74
April	Weekday	19873.98	9300	10573.98

May	Weekday	22958.89	10100	12858.89
May	Weekend	19333.95	9100	10233.95
June	Weekend	16527.14	8400	8127.14
June	Weekday	23400.85	8700	14700.85
July	Weekday	12479.22	8300	4179.22
July	Weekend	17028.18	7700	9328.18
August	Weekday	15653.81	7900	7753.81
August	Weekend	14487.65	8000	6487.65
September	Weekend	12171.60	9200	2971.60
September	Weekday	20178.33	9500	10678.33
October	Weekend	13136.03	7500	5636.03
October	Weekday	18097.12	8800	9297.12
November	Weekday	7787.17	7000	787.17
November	Weekend	5121.42	7200	-2078.58
December	Weekday	7132.78	9800	-2667.22
December	Weekend	6666.80	6700	-33.20

The primary purpose of conducting simulations in our revenue management strategy was to accurately predict and optimize the financial outcomes under various booking scenarios. By employing gamma distribution models, we were able to mimic the real-world randomness and variability of guest arrivals and room bookings. This approach enabled us to test the efficacy of our pricing strategies and room allocation policies under controlled yet realistic conditions.

For each scenario, we ran 1,000 iterations, which allowed us to capture a broad spectrum of possible outcomes and provided a robust statistical base to calculate the expected profits and their variability. In each iteration, the number of rooms booked was generated from the gamma distribution, and the revenue was then calculated based on the adjusted daily rates, which consider both group and transient rates along with transient protection.



The variability in the simulation results underscored the necessity of a dynamic and flexible approach to room pricing and allocation. By understanding which months and day types are prone to higher variability, we can tailor our strategies to either capitalize on high-demand periods or mitigate risks during lower-demand times.

- Peak Season Management: The simulations revealed that months like June and July on weekends have higher expected profits with lower variability, indicating strong and consistent demand. This insight supports a strategy of premium pricing during these times.
- Off-Peak Adjustments: Conversely, the lower expected profits and higher variability in months like November and December suggest a need for more aggressive promotions or discounts to boost occupancy rates.

Overall, the simulation process validated our use of advanced analytical models in setting room rates and allocating room availability. It provided a scientific basis for our decision-making, allowing us to enhance revenue predictably and manage risk effectively.

CONCLUSION

Our team of aspiring data analysts and scientists successfully conducted an in-depth study on revenue management for a 100-room hotel, utilizing the "Hotel Booking Demand" dataset. We are immensely proud of our work, which not only honed our analytical skills but also provided profound insights into the practical aspects of data-driven decision making in the hospitality industry.

Through this project, we developed a sophisticated capacity control model that effectively balances room allocations between group and transient bookings, optimizing for maximum revenue. This model allowed us to dynamically adjust room allocations based on predictive analytics, which included the implementation of gamma distribution functions to forecast booking trends accurately. Our approach led to a deeper understanding of how statistical models can guide strategic decisions in room pricing and allocation to enhance profitability.

The insights gained from this project are substantial:

- Enhanced Analytical Skills: The project provided us with a hands-on opportunity to apply complex statistical methods and predictive modeling in a real-world setting, enhancing our capabilities in data manipulation, analysis, and interpretation.

- Strategic Decision Making: By integrating data analytics with revenue management strategies, we learned how to make informed decisions that align with business objectives, such as maximizing occupancy and revenue.
- Practical Application: The application of theories and models in a real dataset allowed us to see the tangible impacts of our work, improving our confidence and competence in handling large-scale data projects.

Additionally, the project reinforced the importance of clean and relevant data in generating accurate and useful insights. Filtering the dataset to focus on non-canceled bookings enabled us to derive conclusions that accurately reflected true customer behavior and hotel performance.

Looking ahead, we are excited about the potential applications of our findings. The models and strategies developed can be integrated into live hotel booking systems for real-time adjustments, expanded across different properties, or further refined through the adoption of more advanced machine learning techniques. Each of these pathways offers opportunities for continued learning and professional growth.

In summary, this project was a pivotal step in our journey as data scientists. It underscored the vital role of data analytics in strategic business decision-making and equipped us with the tools and knowledge to contribute meaningfully to the field of hospitality management and beyond. We are enthusiastic about bringing these insights into future roles and continuing to develop our skills in this dynamic and impactful arena.